Check for updates

OPEN ACCESS

EDITED BY Shuo Chen, Northeastern University, China

REVIEWED BY Kausik Basak, JIS Institute of Advanced Studies and Research, India Shuang Chang, Vanderbilt University, United States

*CORRESPONDENCE Cléophace Akitegetse, Iso cleo@ziliahealth.com

Elahe Parham, elahe.parham@cervo.ulaval.ca

RECEIVED 27 September 2024 ACCEPTED 01 November 2024 PUBLISHED 14 November 2024

CITATION

Parham E, Munro J, Lapointe N, Landry P, Robidoux J, Brouard D, Quémener M, Parent M, Sauvageau D, Côté DC and Akitegetse C (2024) Estimating retinal blood oxygenation from diffuse reflectance spectra of semi-infinite tissue using principal component analysis. *Front. Photonics* 5:1502799. doi: 10.3389/fphot.2024.1502799

COPYRIGHT

© 2024 Parham, Munro, Lapointe, Landry, Robidoux, Brouard, Quémener, Parent, Sauvageau, Côté and Akitegetse. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Estimating retinal blood oxygenation from diffuse reflectance spectra of semi-infinite tissue using principal component analysis

Elahe Parham^{1,2}*, Jonathan Munro², Nicolas Lapointe³, Patricia Landry⁴, Jonathan Robidoux⁴, Danny Brouard⁴, Mireille Quémener^{1,2}, Martin Parent², Dominic Sauvageau^{3,5}, Daniel C. Côté^{1,2} and Cléophace Akitegetse³*

¹Department of Physics, Physics Engineering and Optics, Université Laval, Québec, QC, Canada, ²CERVO Brain Research, Québec, QC, Canada, ³Zilia INC., Québec, QC, Canada, ⁴Affaires Médicales et Innovation, Héma-Québec, Québec, QC, Canada, ⁵Chemical and Materials Engineering, University of Alberta, Edmonton, AB, Canada

Diffuse reflectance spectroscopy (DRS) is a promising technique for non-invasive monitoring of tissue oxygen saturation (StO₂). However, the interpretation of DRS data can be complicated by the presence of confounding factors such as the volume fraction of blood, tissue scattering, and lipid content which both absorb and scatter. Principal component analysis (PCA) is a multivariate statistical method that can help overcome these challenges by extracting relevant information from complex datasets and providing new dimensions used to estimate parameters such as concentrations. In this study, we present a PCA-based algorithm for estimating retinal StO₂ from DRS measurements. We evaluated the performance of our algorithm using simulated data and experimental measurements on a retinal tissue phantom model. Our results show that the PCA-based algorithm can estimate the value of StO₂ with a root-mean-square error of 6.38% in the presence of confounding factors. Our study demonstrates the potential of PCA as a powerful tool for extracting the concentration of components from complex DRS.

KEYWORDS

diffuse reflectance spectroscopy (DRS), tissue oxygen saturation (StO_2), principal component analysis (PCA), retinal oxygenation, retinal tissue phantom

1 Introduction

Blood oxygen saturation (SO₂ when evaluated in blood vessels, StO₂ when evaluated in tissues) is a vital physiological measure that quantifies the proportion of oxygen-bound hemoglobin in the bloodstream relative to the total available hemoglobin molecules (Hafen and Sharma, 2018). This parameter plays a crucial role in tissue physiology (Swartz and Dunn, 2003) and serves as a diagnostic indicator for a variety of conditions, including heart, cerebral, and vascular diseases (Hoke et al., 2002; Zhang et al., 2011; McCully et al., 1994). Recent advancements have facilitated the measurement of oxygen saturation in the retina, which can be used to detect pathological changes and diseases such as diabetic retinopathy

10.3389/fphot.2024.1502799

and age-related macular degeneration (Linsenmeier and Zhang, 2017). Abnormal SO_2 has been associated with a range of health issues including cancer, diabetes, choroidal disorders, and cardiovascular health (Carmeliet and Jain, 2000; Wittenberg and Wittenberg, 1989; Heitmar and Blann, 2022). Numerous research studies have indicated that irregularities in oxygen levels in the eye fundus are linked to various eye diseases including glaucoma (Boeckaert et al., 2012; Olafsdottir et al., 2014; Tobe et al., 2013), diabetic retinopathy (Jørgensen et al., 2014; Hardarson and Stefánsson, 2012; Guduru et al., 2016), retinal vessel blockages (Williamson et al., 2009; Yoneya et al., 2002), retinitis pigmentosa (Battu et al., 2015), and age-related macular degeneration (Geirsdottir et al., 2014). In addition, retinal oximetry, in which blood oxygenation is evaluated in the large retinal blood vessels, has proven to be effective in non-invasively monitoring certain neurodegenerative diseases (Stefánsson et al., 2017).

Diffuse reflectance spectroscopy (DRS) stands as a potent tool in this context and can provide molecular-level information on tissue, including evaluating StO₂ based on the reversible binding of oxygen to hemoglobin which changes its absorption spectrum (Stokes, 1864). The diffuse reflectance spectrum, obtained using an illumination light source that is commonly UV- visible and can extend to the near-infrared (NIR), is collected using a spectrometer (Delori, 1988; Sircan-Kucuksayan et al., 2015). The absorption coefficient (μ_a), the reduced scattering coefficient (μ_s), and the light path length into the tissue are used to calculate the absorber concentration in a semi-infinite medium according to the modified Beer-Lambert law.

DRS has been widely applied non-invasively in biological tissues for various purposes, primarily to quantify the concentrations of chromophores such as melanin and hemoglobin (Malin et al., 1999; Bender et al., 2009; Troncoso et al., 2021; Ridder et al., 2005; Hani et al., 2011; Hernández et al., 2009). Among these, oxygenated and deoxygenated hemoglobin (oxyhemoglobin and deoxyhemoglobin, respectively) are particularly significant, as they provide critical information on StO₂ and blood volume fraction (BVF). Both parameters are essential for disease diagnosis, tissue health assessment, and cancer monitoring (Chen and Lin, 2010; Claridge et al., 2007). BVF, representing the proportion of tissue occupied by blood (Claridge et al., 2007), offers valuable insights, particularly in understanding tumor microcirculation, predicting tumor behavior, and evaluating treatment responses (Qi et al., 2008). BVF varies according to tissue type and region of interest; for instance, it ranges from 1% to 10% in colon samples (Claridge et al., 2007), 0.2%-7% in skin (Yudovsky and Pilon, 2010), and averages around 8.4% in liver tissue (Chen et al., 2011). Techniques used to measure blood oxygen saturation in the eye fundus often rely on analyzing two or more wavelengths (Hickam et al., 1963; Pittman and Duling, 1975), with the selection of these wavelengths influencing the distinction between blood vessels and surrounding tissue. These approaches typically require userspecific calibration to optimize their accuracy and performance (Delori, 1988; Harris et al., 2003; Schweitzer et al., 1999; Smith, 1999).

Despite the many algorithms developed for this purpose, achieving accurate *in vivo* measurements of absorber concentration in semi-infinite media, such as biological tissues, remains challenging. The presence of lipids in the tissue can have a significant effect on its absorption and scattering properties and alter the intensity and shape of the diffuse reflectance spectrum, especially in the visible and NIR range. Calculating StO_2 from the diffuse reflectance spectrum is also complicated by the sample geometry, such as thickness or curvature, or the composition of the tissue being assessed.

One approach is to simplify all the equations by ignoring the contribution of the non-blood layer in the tissue. Therefore, the computational complexity is reduced. However, the accuracy is decreased by considering that the reflected light would only travel from the blood layer and neglecting the effect of the tissue pigment (Delori and Pflibsen, 1989; Rajaram et al., 2010). If we incorporate the optical properties of the non-blood components, this method can provide a more comprehensive explanation of the tissue, yet there needs to be prior knowledge of the concentration of the non-blood part of the tissue (Cui et al., 1990). Many algorithms are mostly based on parameters depending on the optical setup and numerical computations of coefficients (Middelburg et al., 2011; Reif et al., 2008; Akitegetse et al., 2022).

Here, we propose a principal component analysis (PCA)-based algorithm to evaluate the blood oxygen saturation in retinal tissue using DRS. PCA is used to find independent bases associated with the variations in the dataset. Consequently, the oxygen saturation is calculated by the projection of the dataset in the independent bases onto the spectrum of the compounds of the dataset. The method's performance was evaluated by using both simulated data and data collected on a liquid phantom mimicking retinal properties, showing consistent performance on both datasets. This novel approach is poised to improve the assessment of blood oxygen saturation in the eye and in other tissues. It is expected to strengthen our grasp of a wide range of health conditions and facilitate the development of more precise and comprehensive diagnostic procedures.

2 Method

2.1 Simulated data

Simulated data were created using modified Beer-Lambert law to obtain absorption coefficient spectra similar to Equation 1 for a tissue.

$$A(\lambda) = C_{oxy}\mu_{oxy}(\lambda) + C_{deoxy}\mu_{deoxy}(\lambda) + C_m\mu_m(\lambda) + C_S\log(1/\lambda)$$
(1)

Where C_{oxy} , C_{deoxy} , and C_m are the concentrations of the main photosorbent compounds present in retinal tissue, oxyhemoglobin, deoxyhemoglobin, and melanin, and μ_{oxy} , μ_{deoxy} , μ_m are their respective molar extinction coefficients. The term $C_S \log(1/\lambda)$ accounts for the scattering logarithmic decay due to the effect of blood cells in the tissue being assessed (Schweitzer et al., 2001). The values for the different parameters were taken from the literature (Budnik, 2017; Jacques, 2018; Kessel et al., 2010). These parameters were used to create two sets of data to study the precision of the estimated StO₂ values. In each dataset, the scattering coefficient was uniformly sampled from a range of 0.001–1.

A first dataset was prepared to study the effect of the melanin content on the calculated StO_2 values. We performed simulations

with 50 mmol/L, 100 mmol/L, 200 mmol/L, 300 mmol/L, and 400 mmol/L melanin content (1000 spectra for each melanin content). Additionally, spectra were generated for StO_2 values ranging from 50% to 95% with a uniform random StO_2 distribution and the BVF ranging from 1% to 7.5%. This approach allows for modeling variations in local vessel and capillary density in both healthy and diseased tissues.

A second dataset was employed to evaluate the precision of the algorithm when the BVF had different values. 1000 spectra were simulated for each BVF value (0.75%, 1.5%, 2%, 4%, and 6%) with uniform random StO₂ values ranging from 50% to 95% and uniform random melanin content ranging from 1 mmol/L to 500 mmol/L.

The third dataset was developed to evaluate the performance of the proposed algorithm compared to three other algorithms from literature: the 4-wavelength method (Hammer et al., 2002), the 3wavelength method (Middelburg et al., 2011), and the non-negative least-square (NNLS) method (Akitegetse et al., 2022). The simulated dataset included 5000 spectra with uniform random distribution values for the concentration of different components. The StO₂ ranged between 50% and 95% and BVF varied from 1% to 7.5%. For melanin, the concentrations ranged from 1 mmol/L to 400 mmol/L.

The fourth dataset was developed to be used as the training set for estimating the StO_2 values from phantom samples with different intralipid contents. The simulated dataset included 5000 spectra with uniform random distribution values for the concentration of different components. The StO_2 ranged between 20% and 95% and BVF varied from 4% to 7.5%. For melanin, the concentrations ranged from 1 mmol/L to 400 mmol/L.

The coefficients for oxyhemoglobin and deoxyhemoglobin are composed of the percentage of the oxy- and deoxyhemoglobin and the value of BVF. The summation of all the concentrations of hemoglobin equaled 1.0. In DRS, photon noise is a key source of variability, arising from the inherent randomness in photon detection, which follows a Poisson distribution. As the number of detected photons increases, the signal fluctuations scale with the square root of the photon count. This type of noise predominates when the signal is high, surpassing other noise sources such as dark current or read noise. The photon noise was considered in the simulations by adding a random number generated with Poisson distribution to the intensity of the raw reference spectrum at each wavelength.

2.2 Experimental tissue phantom data

The tissue phantom model used in this study was a previously established retina liquid phantom, comprising varying concentrations of hemoglobin and intralipids (Akitegetse et al., 2022). It was made of a waterproof black chamber and a VIS-NIR-coated achromatic lens with a 17.5-mm focal length inserted through an iris of 8-mm diameter. A 7-mL test tube was inserted into the chamber, at 22 mm from the lens. Distilled water was used in the chamber to simulate the vitreous and the test tube was filled with a mixture of phosphate-buffered saline (PBS, pH 7.4), human blood from human erythrocyte concentrates and lipid emulsion (intralipid, 20% w/v).

Approval for this study was granted by the research ethics committee at Héma-Québec, the blood operator in the province of

Québec, Canada. Healthy volunteers provided whole blood donations using the Leukotrap RC system (Haemonetics Corp., Braintree, MA), yielding approximately 485 mL (\pm 10%) of blood. Within 24 h, the blood was centrifuged, and components were separated using the MacoPress automated blood components separator (Macopharma, France). Red blood cell concentrates (RCC) units underwent leukoreduction and immediate deoxygenation with the Hemanext One device (Hemanext Inc.) for 3 h at room temperature under constant agitation before storage at 2–6°C. Donation processing ensures the availability of red blood cells and preservation for potential use in experiments across different days.

Spectra were obtained using a spectrometer (Hamamatsu) with a sensitivity range from 330 nm to 835 nm with a resolution of 1 nm. The illumination source was a white light-emitting diode ranging from 495 nm to 750 nm.

The actual oxygen saturation of the samples was assessed using a blood gas analyzer (ABL90 Flex Plus, Radiometer). The range of oxygen saturation tested was 20%–100%. Further details on the estimation of the oxygen saturation of the samples can be found in Akitegetse et al., 2022 (Akitegetse et al., 2022).

2.3 Principal components analysis

The absorbance spectra, $A(\lambda)$, of the DRS data were calculated using the following (Equation 2),

$$A(\lambda) = -log\left(\frac{I_{raw}(\lambda) - I_{back}(\lambda)}{I_{ref}(\lambda)}\right)$$
(2)

where I_{raw} is the reflection spectrum, I_{back} is the baseline spectrum, and I_{ref} is the spectrum of the illumination light.

The spectra thus obtained were cropped to retain only the spectral range between 530 nm and 585 nm. This specific wavelength region is crucial because it encompasses the characteristic peaks associated with oxyhemoglobin and deoxyhemoglobin, which are of particular significance in our analysis (Akitegetse et al., 2023). The presence and abundance of these components in the tissue yield the absorbance spectrum of a sample, which can be written using the modified Beer-Lambert law described in Equation 1.

We applied principal component analysis (PCA) for dimensionality reduction, focusing on the first four principal components (PCs) that capture the primary sources of variance associated with oxyhemoglobin, deoxyhemoglobin, melanin, and lipid. These PCs are expected to contain the critical information required for estimating the concentrations of these compounds. Our objective was to estimate the concentrations using a PCA-based algorithm that leverages a change of basis approach (Greenberg, 1998). To implement PCA, we projected the spectra onto the eigenvector basis, where they were treated as independent and reflective of the variations within the dataset. Given that we anticipate four primary compounds in the retina, we conducted the analysis using four independent bases. We assumed V = $\{e_1, e_2, e_3, e_4\}$ to be the basis for PCA decomposition and B = $\left\{e_{oxy}^{\prime}, e_{deoxy}^{\prime}, e_{m}^{\prime}, e_{s}^{\prime}\right\}$ to be the absorption spectra basis of the main components (oxyhemoglobin, deoxyhemoglobin, melanin,



and scattering) in the tissue. A given absorption spectrum A could then be written in both spaces as

$$A = c_1 e_1 + c_2 e_2 + c_3 e_3 + c_4 e_4 \tag{3}$$

$$A = c_{oxy}e'_{oxy} + c_{deoxy}e'_{deoxy} + c_me'_m + c_se'_s$$
(4)

Equation 3 uses the coefficients of the dataset in the new space of the PCA basis and Equation 4 shows the combination of the compounds at different concentrations. To find the concentration of the components c_i (j = oxy, deoxy, m, s), we expanded e_i in terms of *B* as

$$e_{i} = q_{i,oxy} e_{oxy} + q_{i,deoxy} e_{deoxy} + q_{1,m} e_{m} + q_{i,s} e_{s}, i = 1, ..., 4$$
(5)

From Equation 5, the transformation matrix from V to B $[q_{i,j}]$ (i = 1, ..., 4, j = oxy, deoxy, m, s) could be calculated using Equation 6.

$$q_{i,i} = e_i \cdot e_j \tag{6}$$

Therefore, the spectrum obtained from the retina could be rewritten in space B by replacing Equation 5 in Equation 3 as

$$A = \left(c_1 q_{1,oxy} + \dots + c_4 q_{4,s}\right) e'_{oxy} + \dots + \left(c_1 q_{n,s} + \dots + c_4 q_{4,s}\right) e'_s \quad (7)$$

By comparing Equation 7 with Equation 4, the values of c_{oxy} and c_{deoxy} in the assessed retina sample were determined. Finally, the concentrations of oxyhemoglobin and deoxyhemoglobin are used to calculate the StO_2 based on Equation 8.

$$StO_2 = \frac{c_{oxy}}{\left(c_{oxy} + c_{deoxy}\right)} \cdot 100\%$$
(8)

The flowchart of the proposed PCA-based algorithm is shown in Figure 1. The program was written in Python v3.9. The calculated

absorption spectra were cropped as described above and PCA (python, sklearn package) was performed to obtain an appropriate independent basis. The absorbance coefficient of each tissue component was transformed into the same PCA basis. The dot product of the absorption spectra in the new space and the inverse of the transformed components yielded the concentration of each component.

The simulated dataset described in Section 2.1 was used to train the PCA model to evaluate the performance of the PCA-based algorithm. It should be noted that, when employing a simulated dataset, the PCA training utilized 90% of the data, with the remaining 10% dedicated to evaluating the algorithm's ability to predict StO_2 values not encountered during PCA training.

The performance of the algorithm was then evaluated using experimental data from the tissue phantom containing blood at different oxygen saturation levels. Additionally, the tissue phantom data obtained with different hemoglobin concentrations and lens yellowing were used to train a second PCA model and test the performance of the proposed algorithm. While the algorithm demonstrates satisfactory performance on simulated data, reliance on simulations prior to StO2 concentration estimation may introduce an unnecessary step. Consequently, validating the performance of the algorithm on phantom data and concurrently estimating StO_2 concentrations offered a more direct approach, thereby eliminating the need for simulation.

2.4 Comparison of algorithms

To evaluate the performance of the proposed algorithm, we compared the performance of the PCA-derived algorithm to three other algorithms. The first one was the 4-wavelength method



introduced by Hammer et al. (Hammer et al., 2002). This algorithm was developed to compensate for the effect of non-hemoglobin absorption as well as tissue scattering on blood spectra. It uses the absorbance spectra of oxyhemoglobin and deoxyhemoglobin as references at 522 nm, 560 nm, and 569 nm as isosbestic points, and the measured spectrum is transformed between the absorbance of oxyhemoglobin and deoxyhemoglobin spectra. The value of the measured spectrum at 560 nm is then compared to the values of oxyhemoglobin and deoxyhemoglobin at 560 nm to calculate oxygen saturation. The second algorithm for comparison was a three-wavelength method using two isosbestic wavelengths (569 and 586 nm) and one non-isosbestic wavelength (656 nm) (Delori, 1988; Akitegetse et al., 2023). The third algorithm for comparison was a linear regression that was used with a NNLS (Virtanen et al., 2020). To better compare these multi-wavelength algorithms with the PCA-based algorithm, we have used the same wavelength range (from 530 nm to 585 nm) and the same components, namely, oxyhemoglobin, deoxyhemoglobin, melanin, and the scattering term (Akitegetse et al., 2022).

3 Results

To evaluate the performance of the developed PCA-based algorithm, StO_2 values were calculated on the simulated dataset and the root-mean-square error (RMSE) was calculated for each condition tested and for each algorithm tested. The average value of

RMSE and the standard deviation are shown in Figure 2. Figure 2A shows that the RMSE values for all simulations with different melanin content evaluated with the PCA-based algorithm were under 1% while the estimated values were highly correlated with the true values in all cases (r-value >0.99, *p*-value <0.05). Figure 2B demonstrates the RMSE for different BVF values between the calculated StO₂ and the true values. All calculated values in each simulation were highly correlated with the real values (r-value >0.99, *p*-value <0.05). The error in StO₂ values decreased with a higher content of blood. Additionally, the algorithm was robust to different blood volume fractions or melanin contents, since in all cases the standard deviation of the RMSE was less than 0.15%.

Afterward, the model's performance was compared against three alternative methods—namely, the 4-wavelength method, the 3-wavelength method, and the NNLS method. We generated 5000 simulated spectra five times. The RMSE values were computed and depicted in Figure 2C for each dataset. For the 4-wavelength method, the RMSE was approximately 30%, with a Pearson correlation coefficient of 0.69 between estimated StO2 values and actual values. In contrast, the three-wavelength algorithm showed an average RMSE of 12.9% and a higher Pearson correlation coefficient of 0.83 with the actual StO2 values. Furthermore, the NNLS algorithm yielded an average RMSE of 3.85% with a correlation coefficient of 0.97, while the PCA-based algorithm demonstrated the lowest RMSE at 1.68% and the highest correlation coefficient of 0.99 among all algorithms tested. These results indicate that the PCA-based algorithm achieves the highest accuracy compared to the other algorithms assessed.



FIGURE 3

Performance of PCA-based algorithm. (A) Calculated StO_2 using the PCA-based algorithm compared to ground truth using 900 simulated spectra for training and 100 simulated spectra for testing. (B) Calculated StO_2 in the tissue phantom model with different oxygen saturation levels and intralipids (IL) concentrations. Five acquisitions (n = 5) were done from each sample and the mean \pm standard deviation is displayed.



The performance of the PCA-based algorithm was evaluated on the fourth simulated dataset. Using the same fitted PCA on the simulated data, the StO2 values of the phantom data with different intralipid concentrations were estimated. The coefficient of determination for the calculated oxygen saturation with the PCAbased algorithm on the simulated data (shown in Figure 3A) was 0.98 and the overall RMSE was 2.78%. The results of using the same fitted PCA on the data from the tissue phantom model with various intralipid concentrations are shown in Figure 3B. Samples at different StO₂ values (29.5%, 56.1%, 72.3%, and 91.9%) and intralipid contents (0.25%, 0.72%, 0.95%, and 1.80% IL) were assessed through five acquisitions each (49 spectra per acquisition). Figure 3 shows the calculated StO2 values for tissue phantom blood samples at different StO₂ levels and lipid contents. The average RMSE between the StO₂ values of each acquisition and the real values of all acquisitions made with the PCA-based algorithm was 7.38%.

To evaluate the performance of the algorithm with tissues having different micro-vessels densities, samples at five different concentrations of hemoglobin (0.25 g/dL, 0.50 g/dL, 1.00 g/dL, 2.25 g/dL, and 4.50 g/dL) were prepared and assessed in the tissue phantom. A total of 100 acquisitions were performed with samples at different StO2 values. The absolute error between the calculated StO₂ and the value obtained from a blood gas analyzer is shown in Figure 4 for the range of hemoglobin concentrations. The average error, accounting for all cases tested, was less than 6.3%, and we found there was no correlation between the error on the StO₂ values and the concentration of hemoglobin (r = 0.69, p-value = 0.19).

The effect of lens yellowing on the performance of the algorithm is shown in Figure 5. There were 15 samples for each level of StO_2 , with three different levels of yellowing of the lens (without yellowing, Y20, and Y60). The concentration of hemoglobin and intralipid content were constant (at 1 g/dL and 0.25%, respectively) in the samples. The statistical equivalency of the values obtained for each of the three groups of measurements showed the algorithm is robust to the yellowing of the lens.

4 Discussion

A PCA-based algorithm to estimate the blood oxygen saturation in a semi-infinite tissue, specifically micro-vascularized regions of the retina, using DRS has been introduced. The error on StO_2 calculated by the algorithm was lower than that of previously established algorithms based on four-wavelength, threewavelength, and NNLS analyses. The accuracy of values measured by 2-, 3- or 4-wavelength methods can be affected by artifacts, signal-to-noise ratio, optical interference, and the accuracy of the oxygen standard that is used for calibration (Sinex, 1999). The 4-wavelength method uses absorbance values at four different wavelengths to estimate SO_2 or StO_2 , three of them are for transforming the measured spectrum and only one wavelength is



used for quantification (Pittman and Duling, 1975). Hence, this approach is vulnerable to noise, especially at 560 nm. While multiwavelength techniques optimize regressors, they can result in overfitting or underfitting when considering missing or excessive components in regression, leading to less accurate estimates of oxygen saturation (Su et al., 2012).

The performance of the PCA algorithm on the simulated dataset with different concentrations of melanin shows a slight decrease in accuracy as the melanin content increases. It is important to note that the relationship between melanin content and the accuracy of the algorithm may vary depending on the specific algorithm. Melanin is a pigment responsible for absorbing light in the visible and NIR spectrum. When the melanin content is higher, it absorbs more light, leading to a decrease in the intensity of light reaching deeper tissue layers and returning from the eye. Reducing the fraction of captured light that has undergone interaction with hemoglobin typically results in the attenuation of the characteristic peaks observed in hemoglobin (Damodaran et al., 2018).

Based on the results of our simulations, the calculated errors for different BVF values demonstrate a slight increase in the accuracy of StO_2 estimation (less than 1%) as the volume of blood increases. This can happen for several reasons. When blood volume is greater, there is a greater amount of hemoglobin available to absorb light which provides more robust spectral information, making it easier for the algorithm to estimate oxygen saturation accurately. This causes the backscattered signal from blood cells to become stronger compared to the background tissue. This improved signal-to-noise ratio makes it easier for the algorithm to extract the oxygen saturation information from the spectral measurements, leading to increased accuracy.

The performance of the algorithm on the data acquired from samples placed in tissue phantoms, using different levels of intralipid and hemoglobin concentration, is evaluated in systems with different optical properties including scattering and absorption. This approach allows us to assess how well the performance of the algorithm under varying conditions and provides a benchmark for understanding the algorithm's limitations and strengths. Moreover, phantoms provide controlled experimental conditions where the actual oxygen saturation values can be precisely determined using a blood gas analyzer, the gold standard in the field. By comparing the algorithm's measurements with known values, researchers can validate its accuracy and identify any discrepancies or biases that need to be addressed.

The PCA-based algorithm showed that when the volume of blood is greater, we observe consistent results when the intralipid and hemoglobin contents are changing. Unlike simulations, experiments conducted on phantoms do not exhibit a decrease in error as the hemoglobin content increases, and it is confirmed by the literature that the blood volume does not affect the accuracy of SO₂ estimation (Chen et al., 2015).

Measurements made from simulated spectra showed superior performance than those from phantom experiments, with errors smaller than 1%. In addition, in phantom experiments, the error increased with greater intralipid contents. This difference in performance can be attributed to several factors. Simulations provide a highly controlled environment in which scattering and absorption are assumed to be homogeneous, whereas phantom experiments introduce variability due to factors such as uneven intralipid mixing, surface scattering effects, and the optical heterogeneity of the phantom material. These complexities make light-tissue interactions in phantom models more intricate than in simulations. Moreover, the algorithm used to estimate oxygen saturation assumes a homogeneous scattering and absorption environment, which is an oversimplification in simulations. In contrast, phantom data exhibit more variable scattering properties that may deviate from the idealized model, thus reducing algorithm accuracy. Consequently, phantom data serve as a more realistic proxy for real tissues, offering a more representative assessment of performance in *in vivo* applications.

Finally, the performance of the algorithm was evaluated using lenses that simulate various degrees of yellowness. Yellowing of the eye significantly alters the illumination of the fundus by filtering the blue and green light. Comparing the results of two lenses with different yellowing levels and a lens without yellowness demonstrates that the algorithm remains robust when the blue and green light are filtered out. These promising results are a positive step forward toward assessing StO2 in patients with cataracts, although it's worth acknowledging that they do not provide an absolute guarantee of the algorithm's robustness against all types of cataracts. In particular, cortical cataracts, which not only alter the light spectrum but also cause clouding of the lens, are expected to pose significant challenges. The reduced amount of light reaching and returning from the eye's fundus can lower sensitivity and introduce more errors in StO2 measurements. In this scenario, the problem extends beyond merely adjusting the oximetry algorithm-it also involves identifying a more effective method to allow sufficient light transmission and collection to ensure accurate measurements.

Our evaluation of the algorithm's performance involved two distinct approaches. In the first approach, depicted in Figures 2, 3, we utilized simulated data for training PCA, subsequently evaluating its performance on novel data not encountered during training. Conversely, for the outcomes presented in Figures 4, 5, we employed phantom data for both training PCA and assessing algorithmic performance. Remarkably, the algorithm exhibited robust performance in both scenarios, underscoring its versatility. PCA, in this context, is not employed for conventional feature extraction; rather, it serves to establish a new basis for spectral analysis. Consequently, the algorithm's efficacy is upheld when trained with a dataset showcasing variations in the concentration of a target component, whether simulated or experimental (phantom).

The PCA-based method introduced here has the benefit of having no fit and weights to estimate, so it could be run on whole datasets without the need to exclude noisy spectra. The method is easy to implement and there is less computational complexity compared to methods with numerical or networkbased AI analysis. More importantly, using PCA in the algorithm mitigates overfitting issues. To further prevent overfitting or underfitting, we ensured that the dataset included sufficient variability to represent real-world scenarios. Yet the algorithm needs to be validated with the results from real eyes since the simulations and the phantom do not perfectly represent a real tissue; however, this is a complex endeavor as there is no gold standard to establish a ground truth for blood oxygen saturation in the tissues of the retina.

The tests conducted with tissue phantoms allowed us to evaluate the strength and accuracy of our oximetry algorithm, even in the presence of interfering elements. While the phantom incorporated key structural features of the eye, it remains a simplified representation of its actual complexity. This model provided a controlled environment to isolate specific factors, facilitating reallife mimicking measurements and validating the data collection process. However, it does not account for critical variables such as varying thickness, optical properties, or the composition of different layers in the eye's fundus. Furthermore, using red blood cells in saline solution combined with components like intralipids presents limitations in the modeling of real tissues. The assumption of homogeneous tissue properties made in both phantom and simulation models overlooks the influence of factors such as variable blood flow velocities and red blood cell aggregation in capillaries, which can significantly affect the tissue's optical properties, including light absorption and scattering, and potentially impact StO₂ measurements. Future studies should focus on developing more complex phantoms that better mimic the eye and incorporate realistic parameters, as well as simulating blood flow with whole blood and employing advanced models to accurately represent tissue behavior.

5 Conclusion

Our study introduces a PCA-based algorithm for estimating StO_2 using DRS in the retina. The algorithm demonstrates superior performance compared to existing methods, such as 4-wavelength, 3-wavelength, and NNLS algorithms, with reduced errors in estimated values. Our evaluations encompassed simulated data, experimental measurements on tissue phantoms, and assessments under various conditions, highlighting the algorithm's robustness and versatility.

The discussion emphasizes the impact of confounding factors, such as melanin content, blood volume fraction, and optical properties of tissues, on the algorithm's accuracy. It is crucial to recognize the limitations of the algorithm, particularly in the presence of increased melanin content and higher intralipid concentrations, where accuracy may diminish. Additionally, the study scrutinizes the algorithm's performance under conditions simulating different degrees of yellowing of the lenses, showcasing its resilience against alterations in the light spectrum.

The comparison between simulated and phantom data reveals nuances in error rates, emphasizing the challenges associated with real-world tissue complexities. The algorithm's efficacy is upheld in both scenarios, demonstrating its adaptability and robustness. Importantly, our technique employs a simplified tissue model, acknowledging its limitations in capturing the full complexity of real tissue. Recommendations for future improvements include incorporating capillary contributions, simulating blood flow, and utilizing more complex tissue models.

Despite these considerations, the PCA-based algorithm offers several advantages, including no need for fitting or weight estimation, ease of implementation, and lower computational complexity compared to alternative methods. However, it is crucial to validate the algorithm using real eye data, as simulations and phantom models may not perfectly represent the intricacies of actual tissue.

In conclusion, our findings underscore the potential of the PCAbased algorithm as a valuable tool for non-invasive monitoring of oxygen saturation in clinical and research settings. With proper optimization, validation, and consideration of its limitations, the PCA-based approach holds promise for advancing the field of DRS and enhancing our ability to accurately estimate tissue oxygen saturation.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Ethics committee at Héma-Québec. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

EP: Conceptualization, Formal Analysis, Investigation, Methodology, Visualization, Writing-original draft. Writing-review and editing. JM: Investigation, Writing-review and editing. NL: Conceptualization, Writing-review and editing. PL: Investigation, Writing-review and editing. JR: Investigation, Writing-review and editing. DB: Supervision, Writing-review and editing. MQ: Supervision, Writing-review and editing. MP: Conceptualization, Supervision, Writing-review and editing. DS: Conceptualization, Supervision, Writing-review and editing. DC: Conceptualization, Supervision, Writing-review and editing. CA: Conceptualization, Investigation, Methodology, Supervision, Writing-original draft, Writing-review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by the NSERC Discovery Grant to DCC (Grant No.

References

Akitegetse, C., Landry, P., Robidoux, J., Lapointe, N., Brouard, D., and Sauvageau, D. (2022). Monte-Carlo simulation and tissue-phantom model for validation of ocular oximetry. *Biomed. Opt. Express* 13, 2929. doi:10.1364/boe.458079

Akitegetse, C., Poirier, J., and Sauvageau, D. (2023). Sensitivity of visible range multiwavelength algorithms for retinal tissue oximetry to acquisition parameters. *Biomed. Opt. Express* 14, 4296–4309. doi:10.1364/boe.495721

Battu, R., Mohan, A., Khanna, A., Kumar, A., and Shetty, R. (2015). Retinal oxygen saturation in retinitis pigmentosa and macular dystrophies in asian-Indian eyes. *Invest Ophthalmol. Vis. Sci.* 56, 2798–2802. doi:10.1167/iovs.14-15993

Bender, J. E., Shang, A. B., Moretti, E. W., Yu, B., Richards, L. M., and Ramanujam, N. (2009). Noninvasive monitoring of tissue hemoglobin using UV-VIS diffuse reflectance spectroscopy: a pilot study. *Opt. Express* 17, 23396–23409. doi:10.1364/oe.17.023396

Boeckaert, J., Vandewalle, E., and Stalmans, I. (2012). Oximetry: recent insights into retinal vasopathies and glaucoma. *Bull. Soc. Belge. Ophtalmol.*, 75–83.

Budnik, B. J. (2017). A liver-kidney biosubsystem. rc.library.uta.edu. Available at: https://rc.library.uta.edu/uta-ir/handle/10106/26787.

Carmeliet, P., and Jain, R. K. (2000). Angiogenesis in cancer and other diseases. Nature 407, 249-257. doi:10.1038/35025220

Chen, P., Fernald, B., and Lin, W. (2011). Estimation of regional hemoglobin concentration in biological tissues using diffuse reflectance spectroscopy with a novel spectral interpretation algorithm. *Phys. Med. Biol.* 56, 3985–4000. doi:10.1088/0031-9155/56/13/015

Chen, P., and Lin, W. (2010) "Determination of *in vivo* blood oxygen saturation and blood volume fraction using diffuse reflectance spectroscopy," in *IFMBE proceedings*. Berlin, Heidelberg: Springer Berlin Heidelberg, 336–339.

Chen, S., Yi, J., Liu, W., Backman, V., and Zhang, H. F. (2015). Monte Carlo investigation of optical coherence tomography retinal oximetry. *IEEE Trans. Biomed. Eng.* 62, 2308–2315. doi:10.1109/tbme.2015.2424689

Claridge, E., Hidović-Rowe, D., Taniere, P., and Ismail, T. (2007). "Quantifying mucosal blood volume fraction from multispectral images of the colon," in *Medical*

RGPIN-2020-06936) and Mitacs accelerate to MP (No. IT28655). Additionally, this research was funded by Zilia Inc.; Héma-Québec; National Research Council of Canada—Industrial Research Assistance Program; Government of Canada Scientific Research and Experimental Development Tax Incentive Program.

Conflict of interest

Authors NL, DS, and CA were employed by Zilia INC.

The authors declare that this study received funding from Zilia Inc. The funder was involved in all the stages of the study including: the study design, collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication.

Generative AI Statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

imaging 2007: physiology, function, and structure from medical images. Editors A. Manduca, and X. P. Hu (SPIE). doi:10.1117/12.709559

Cui, W. J., Ostrander, L. E., and Lee, B. Y. (1990). *In vivo* reflectance of blood and tissue as a function of light wavelength. *IEEE Trans. Biomed. Eng.* 37, 632–639. doi:10. 1109/10.55667

Damodaran, M., Amelink, A., and de Boer, J. F. (2018). Optimal wavelengths for subdiffuse scanning laser oximetry of the human retina. *J. Biomed. Opt.* 23, 1–15. doi:10. 1117/1.jbo.23.8.086003

Delori, F. C. (1988). Noninvasive technique for oximetry of blood in retinal vessels. *Appl. Opt.* 27, 1113–1125. doi:10.1364/ao.27.001113

Delori, F. C., and Pflibsen, K. P. (1989). Spectral reflectance of the human ocular fundus. Appl. Opt. 28, 1061-1077. doi:10.1364/ao.28.001061

Geirsdottir, A., Hardarson, S. H., Olafsdottir, O. B., and Stefánsson, E. (2014). Retinal oxygen metabolism in exudative age-related macular degeneration. *Acta Ophthalmol.* 92, 27–33. doi:10.1111/aos.12294

Greenberg, M. D. (1998). Advanced engineering mathematics. Pearson Education India.

Guduru, A., Martz, T. G., Waters, A., Kshirsagar, A. V., and Garg, S. (2016). Oxygen saturation of retinal vessels in all stages of diabetic retinopathy and correlation to ultrawide field fluorescein angiography. *Invest Ophthalmol. Vis. Sci.* 57, 5278–5284. doi:10. 1167/iovs.16-20190

Hafen, B. B., and Sharma, S. (2018). Oxygen saturation. Available at: https://europepmc.org/article/nbk/nbk525974.

Hammer, M., Thamm, E., and Schweitzer, D. (2002). A simple algorithm for *in vivo* ocular fundus oximetry compensating for non-haemoglobin absorption and scattering. *Phys. Med. Biol.* 47, N233–N238. doi:10.1088/0031-9155/47/17/403

Hani, A. F. M., Nugroho, H., Noor, N. M., Rahim, K. F., and Baba, R. (2011). "A modified beer-lambert model of skin diffuse reflectance for the determination of melanin pigments," in 5th kuala lumpur international conference on biomedical engineering 2011. Springer Berlin Heidelberg, 393–397.

Hardarson, S. H., and Stefánsson, E. (2012). Retinal oxygen saturation is altered in diabetic retinopathy. Br. J. Ophthalmol. 96, 560-563. doi:10.1136/bjophthalmol-2011-300640

Harris, A., Dinn, R. B., Kagemann, L., and Rechtman, E. (2003). A review of methods for human retinal oximetry. *Ophthalmic Surg. Lasers Imaging* 34, 152–164. doi:10.3928/1542-8877-20030301-16

Heitmar, R., and Blann, A. D. (2022). Oxygen saturation in retinal vessels and their correlation with endothelial microparticles in diabetes mellitus and/or cardiovascular disease. *Microvasc. Res.* 142, 104336. doi:10.1016/j.mvr.2022.104336

Hernández, S. E., Rodríguez, V. D., Pérez, J., Martín, F. A., Castellano, M. A., and Gonzalez-Mora, J. L. (2009). Diffuse reflectance spectroscopy characterization of hemoglobin and intralipid solutions: *in vitro* measurements with continuous variation of absorption and scattering. *J. Biomed. Opt.* 14, 034026. doi:10.1117/1.3149864

Hickam, J. B., Frayser, R., and Ross, J. C. (1963). A study of retinal venous blood oxygen saturation in human subjects by photographic means. *Circulation* 27, 375–385. doi:10.1161/01.cir.27.3.375

Hoke, T. R., Donohue, P. K., Bawa, P. K., Mitchell, R. D., Pathak, A., Rowe, P. C., et al. (2002). Oxygen saturation as a screening test for critical congenital heart disease: a preliminary study. *Pediatr. Cardiol.* 23, 403–409. doi:10.1007/s00246-002-1482-8

Jacques, S. (2018). Extinction coefficient of melanin. Available at: https://omlc.org/ spectra/melanin/.

Jørgensen, C. M., Hardarson, S. H., and Bek, T. (2014). The oxygen saturation in retinal vessels from diabetic patients depends on the severity and type of vision-threatening retinopathy. *Acta Ophthalmol.* 92, 34–39. doi:10.1111/aos.12283

Kessel, L., Lundeman, J. H., Herbst, K., Andersen, T. V., and Larsen, M. (2010). Age-related changes in the transmission properties of the human lens and their relevance to circadian entrainment. *J. Cataract. Refract Surg.* 36, 308–312. doi:10.1016/j.jcrs.2009.08.035

Linsenmeier, R. A., and Zhang, H. F. (2017). Retinal oxygen: from animals to humans. Prog. Retin Eye Res. 58, 115–151. doi:10.1016/j.preteyeres.2017.01.003

Malin, S. F., Ruchti, T. L., Blank, T. B., Thennadil, S. N., and Monfre, S. L. (1999). Noninvasive prediction of glucose by near-infrared diffuse reflectance spectroscopy. *Clin. Chem.* 45, 1651–1658. doi:10.1093/clinchem/45.9.1651

McCully, K. K., Halber, C., and Posner, J. D. (1994). Exercise-induced changes in oxygen saturation in the calf muscles of elderly subjects with peripheral vascular disease. *J. Gerontol.* 49, B128–B134. doi:10.1093/geronj/49.3.b128

Middelburg, T. A., Kanick, S. C., de Haas, E. R. M., Sterenborg, H. J. C. M., Amelink, A., Neumann, M. H. A. M., et al. (2011). Monitoring blood volume and saturation using superficial fibre optic reflectance spectroscopy during PDT of actinic keratosis. *J.*4, 721–730. doi:10.1002/jbio.201100053

Olafsdottir, O. B., Vandewalle, E., Abegão Pinto, L., Geirsdottir, A., De Clerck, E., Stalmans, P., et al. (2014). Retinal oxygen metabolism in healthy subjects and glaucoma patients. *Br. J. Ophthalmol.* 98, 329–333. doi:10.1136/bjophthalmol-2013-303162

Pittman, R. N., and Duling, B. R. (1975). A new method for the measurement of percent oxyhemoglobin. J. Appl. Physiol. 38, 315–320. doi:10.1152/jappl.1975.38.2.315

Qi, X.-L., Burns, P., Hong, J., Stainsby, J., and Wright, G. (2008). Characterizing blood volume fraction (BVF) in a VX2 tumor. *Magn. Reson Imaging* 26, 206–214. doi:10.1016/j.mri.2007.05.010

Rajaram, N., Gopal, A., Zhang, X., and Tunnell, J. W. (2010). Experimental validation of the effects of microvasculature pigment packaging on *in vivo* diffuse reflectance spectroscopy. *Lasers Surg. Med.* 42, 680–688. doi:10.1002/lsm.20933

Reif, R., Amorosino, M. S., Calabro, K. W., Singh, S. K., and Bigio, I. J. (2008). Analysis of changes in reflectance measurements on biological tissues subjected to different probe pressures. *J. Biomedi. Optics.* Society of Photo-Optical Instrumentation Engineers 13, 010502–010502. doi:10.1002/lsm.20933

Ridder, T. D., Hendee, S. P., and Brown, C. D. (2005). Noninvasive alcohol testing using diffuse reflectance near-infrared spectroscopy. *Appl. Spectrosc.* 59, 181–189. doi:10.1366/0003702053085098

Schweitzer, D., Hammer, M., Kraft, J., Thamm, E., Königsdörffer, E., and Strobel, J. (1999). *In vivo* measurement of the oxygen saturation of retinal vessels in healthy volunteers. *IEEE Trans. Biomed. Eng.* 46, 1454–1465. doi:10. 1109/10.804573

Schweitzer, D., Thamm, E., Hammer, M., and Kraft, J. (2001). A new method for the measurement of oxygen saturation at the human ocular fundus. *Int. Ophthalmol.* 23, 347–353. doi:10.1023/a:1014458815482

Sinex, J. E. (1999). Pulse oximetry: principles and limitations. Am. J. Emerg. Med. 17, 59–66. doi:10.1016/s0735-6757(99)90019-0

Sircan-Kucuksayan, A., Uyuklu, M., and Canpolat, M. (2015). Diffuse reflectance spectroscopy for the measurement of tissue oxygen saturation. *Physiological* 36, 2461–2469. doi:10.1088/0967-3334/36/12/2461

Smith, M. H. (1999). Optimum wavelength combinations for retinal vessel oximetry. *Appl. Opt.* 38, 258–267. doi:10.1364/ao.38.000258

Stefánsson, E., Olafsdottir, O. B., Einarsdottir, A. B., Eliasdottir, T. S., Eysteinsson, T., Vehmeijer, W., et al. (2017). Retinal oximetry discovers novel biomarkers in retinal and brain diseases. *Invest Ophthalmol. Vis. Sci.* 58, BIO227–BIO233. doi:10.1167/iovs.17-21776

Stokes, G. G.VIII (1864). On the reduction and oxidation of the colouring matter of the blood. *Proc. R. Soc. Lond* 13, 355–364. doi:10.1098/rspl.1863.0080

Su, X., Yan, X., and Tsai, C.-L. (2012). Linear regression. Wiley Interdiscip. Rev. Comput. Stat. 4, 275–294. doi:10.1002/wics.1198

Swartz, H. M., and Dunn, J. F. (2003). Measurements of oxygen in tissues: overview and perspectives on methods. *Adv. Exp. Med. Biol.* 530, 1–12. doi:10.1007/978-1-4615-0075-9_1

Tobe, L. A., Harris, A., Schroeder, A., Gerber, A., Holland, S., Amireskandari, A., et al. (2013). Retinal oxygen saturation and metabolism: how does it pertain to glaucoma? An update on the application of retinal oximetry in glaucoma. *Eur. J. Ophthalmol.* 23, 465–472. doi:10.5301/ejo.5000289

Troncoso, J. R., Diaz, P. M., Lee, D. E., Quick, C. M., and Rajaram, N. (2021). Longitudinal monitoring of tumor response to immune checkpoint inhibitors using noninvasive diffuse reflectance spectroscopy. *Biomed. Opt. Express* 12, 3982–3991. doi:10.1364/boe.426879

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* 17, 261–272. doi:10.1038/s41592-019-0686-2

Williamson, T. H., Grewal, J., Gupta, B., Mokete, B., Lim, M., and Fry, C. H. (2009). Measurement of PO2 during vitrectomy for central retinal vein occlusion, a pilot study. *Graefes Arch. Clin. Exp. Ophthalmol.* 247, 1019–1023. doi:10.1007/s00417-009-1072-z

Wittenberg, B. A., and Wittenberg, J. B. (1989). Transport of oxygen in muscle. Annu. Rev. Physiol. 51, 857–878. doi:10.1146/annurev.physiol.51.1.857

Yoneya, S., Saito, T., Nishiyama, Y., Deguchi, T., Takasu, M., Gil, T., et al. (2002). Retinal oxygen saturation levels in patients with central retinal vein occlusion. *Ophthalmology* 109, 1521–1526. doi:10.1016/s0161-6420(02)01109-0

Yudovsky, D., and Pilon, L. (2010). Rapid and accurate estimation of blood saturation, melanin content, and epidermis thickness from spectral diffuse reflectance. *Appl. Opt.* 49, 1707–1719. doi:10.1364/ao.49.001707

Zhang, K., Zhu, L., and Fan, M. (2011). Oxygen, a key factor regulating cell behavior during neurogenesis and cerebral diseases. *Front. Mol. Neurosci.* 4, 5. doi:10.3389/fnmol.2011.00005